Technical Note: Using Hadamard Symmetry to Probe and Scramble Structure

1. Intuition: a built-in “structure detector”

Imagine you had a transformation that splits any signal into two parts that **should never interact**—by design.

Now imagine passing that signal through a system (a channel, a matrix, a neural layer). If those two parts start interacting, you immediately know:

> something in the system is mixing things it shouldn’t.

That gives you a powerful idea:

- design a split with a **hard structural rule**
- measure how much that rule is violated

The Walsh–Hadamard transform gives exactly this kind of split, and diagonal sign matrices D let you **randomize and probe it in different ways**.

---

2. The core tool: an involution

Start with a matrix H such that:

$H^2 = I$

The normalized Walsh–Hadamard matrix is the standard example.

This gives a natural split of space into two orthogonal parts:

$P_+ = ½(I + H), \quad P_- = ½(I - H)$

Any vector becomes:

$x = x_+ + x_-$ $x_+ = P_+x$ $x_- = P_-x$

Key property:

> In an ideal system, x_+ and x_- **do not interact at all**.

---

3. Systems that respect the split

A matrix A preserves this structure if:

$AH = HA$

This means:

- no mixing between the two subspaces
- the split remains “clean” after applying A

If this condition is violated:

- energy leaks between the two parts
- you get a measurable signal of distortion

---

4. You can measure and fix the violation

Any matrix A can be split into:

- structure-preserving part:

  $A_{good} = ½(A + HAH)$

- structure-violating part:

  $A_{bad} = ½(A - HAH)$

Interpretation:

- A_{good} : keeps the split intact
- A_{bad} : causes mixing

You can even **force the structure**:

$A \mapsto \frac{1}{2}(A+HAH)$

---

5. Now introduce randomness: diagonal sign matrices D

Let:

$D = \mathrm{diag}(\pm 1,\ldots,\pm 1)$

This looks trivial—but it's not.

When viewed through the Hadamard structure, D splits into:

- D_{good} preserves the split
- D_{bad}: mixes the two halves

with:
$D_{good} = \frac{1}{2}(D + HDH)$

$D_{bad} = \frac{1}{2}(D - HDH)$

---

6. What actually happens for random D

For a random choice of signs:

- both parts are **large and comparable in size**
- roughly half the "energy" goes into each

So:

> a random D both **mixes within each subspace** and **mixes across them**

This is the key insight.

---

7. Two types of mixing

It helps to think of the space as two “channels”:

- V_+ (positive eigenspace)
- V_- (negative eigenspace)

Then:

- D_{good} : shuffles data **inside each channel**
- D_{bad} : **cross-wires the channels**

A random D:

does both strongly

---

8. Why this produces good randomization

Even though D is diagonal:

- after applying H, everything becomes dense
- entries behave like sums of many ± contributions

This leads to:

- well-spread energy
- low correlation between components
- behavior similar to random projections

---

9. The key construction

The widely used transform:

$x \mapsto HDx$

works because:

1. D: injects randomness (both within and across subspaces)
2. H: spreads everything globally

Together:

> this mimics a dense random transform at a much lower cost

---

10. Variant: changing the split

You can also form:

$H' = DHD$

This is another involution:

- same structure
- different eigenspaces

So:

> each D gives you a **new way to split the space**

---

11. Practical experiments to try

If you want to explore this numerically:

(a) Measure mixing

- compute $x_+ = P_+x$, $x_- = P_-x$
- apply a matrix A
- measure cross-term measure cross-term $\langle Ax_+, Ax_- \rangle$

(b) Split a random D

- compute D_{good}, D_{bad}
- compare their norms
- observe ≈50/50 split

(c) Visualize energy spread

- apply HD to a sparse vector versus a dense vector
- observe how energy distributes across coordinates

(d) Compare transforms

- compare Hx, HDx, DHDx
- look at variance and correlations

---

12. What to expect

In practice:

- HD behaves much more “randomly” than H alone
- energy spreads evenly across components
- correlations drop significantly

---

13. Conceptual takeaway

There are two complementary mechanisms at work:

- **within-subspace mixing** (from D_{good})
- **cross-subspace mixing** (from D_{bad})

Both are necessary for effective randomization.

---

14. Big picture

A simple identity:

$H^2 = I$

gives you:

- a clean structural split
- a way to detect violations
- a way to enforce constraints
- and, when combined with randomness,
  a fast approximation to fully random transforms

---

15. Final intuition

A useful mental model:

- H: defines two non-interacting channels
- D: both shuffles each channel and cross-connects them
- repeated application: destroys structure and spreads information

So even though everything is built from simple pieces:

> the combination behaves like a surprisingly powerful random system

---